Abstract

Topic Modelling has been widely used in the fields of machine learning, text mining etc. It was proposed to generate statistical models to classify multiple topics in a collection of document, and each topic is represented by distribution of words. But many variants of topic models have been proposed and most of them are based on the concept of bag-of-words and it ignores the association of words for representing topics. Nowadays patterns are used for representing topics, since they have more discriminative power than words for representing multiple topics in a document. A detailed survey of some of the most important methods for topic modelling is presented. A brief comparison among the key techniques is also presented to complete the survey.
Review on Pattern based Document Modelling Techniques

1998.


Index Terms

Computer Science       Pattern Recognition
Keywords

Information Retrieval, Information Filtering, Topic models